AcaStat User Manual

Version 8.2 for Mac and Windows

Table of Contents

Introduction

AcaStat provides students, teachers, researchers, and managers with an inexpensive and easy to use data analysis tool and instructional aid. AcaStat includes a spreadsheet for entering and viewing data, statistical procedures and summary statistics modules, an output log, decision tools, and statistics tutorial. A statistics glossary is also provided and purchased versions of AcaStat include the Applied Statistics Handbook.

Getting Help

Application help can be found in this manual and in the Help module provided with AcaStat. A QuickStart guide is also available at http://www.acastat.com/universal.htm.

In addition to application help, statistical help is provided in four formats:

✓ Statistics tutorial

✓ Glossary of statistical terms and definitions

✓ Applied Statistics Handbook in pdf format for licensed versions

✓ Student Workbook in pdf format

Installation

Download an evaluation version or purchase AcaStat from http://www.acastat.com. The software is downloaded in a self-extracting file. Save the software and double-click on the file to begin the installation process.

Uninstalling AcaStat

For Mac versions, drag the program file to the Trash Bin. For Windows versions, use the Windows Control Panel option ADD/REMOVE programs to uninstall AcaStat. Files that were created or changed after installation of AcaStat must be removed separately by the manual file deletion procedures in Mac and Windows.

Terminology Used in the Manual

When discussing statistical procedures, the manual will refer to two general types of data -- continuous and categorical. The following terms will be used for convenience and simplicity, but there are variations in these themes that will be apparent to those with statistics training. Please see the Statistics Glossary in AcaStat for more information on these and related terms.

Continuous Data – This refers to data where it is reasonable to calculate an average. This type of data is also referred to in statistics as interval/ratio level of measurement or scale data.

Categorical Data – This refers to data where averages would not be meaningful. This type of data is also referred to in statistics as nominal/ordinal level of measurement or discrete data.

AcaStat Startup Screen

AcaStat starts by displaying the Data module and an empty spreadsheet. There are also tabs for the Output Viewer, Tutor, and Decision Tools modules. A control panel along the left side of the application contains statistical procedure controls. Use the Statistical Procedures panel to select and run analyses. The Output Viewer module automatically displays the results for the current analysis and stores all analyses conducted after opening AcaStat. The output from all procedures, including the tutorial, is automatically appended to the output log.
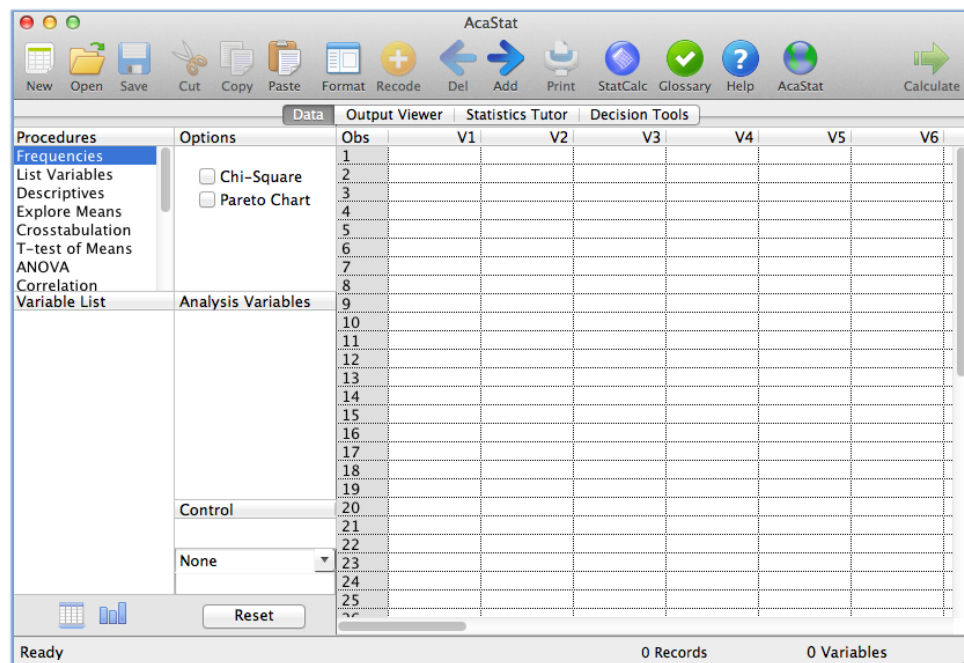
Data Module

The Data module uses a spreadsheet to create, import, and export data files. Opening or importing a new file will replace the current contents.
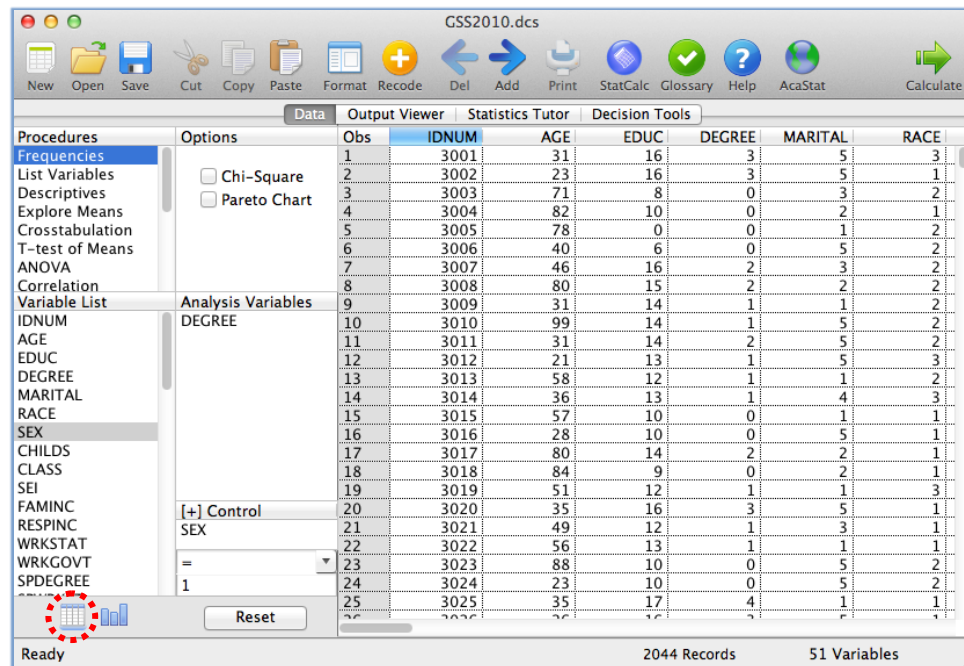
Rows

Each row represents one observation (this is also known as the unit of analysis). If the data contain information on individuals, each row should contain the data for one person (observation). If the data contain information on cities, each row should contain the data relevant to one particular city.

Columns

Each column represents data for all observations relating to one variable or characteristic. As an example, if a column contains information on personal income in U.S. dollars, all data in that column should represent the income for each person in the data file.

Cells

Each cell represents a value for one observation for one variable. It can contain string data (words) or numerical data (numbers). Most research data files use numerical data to represent words. As an example, for the variable sex, it is common to code Male as "0" and Female as

"1." This makes data entry much easier since one number is entered instead of typing a word. It also simplifies recoding variables, identifying controls, and running analyses.

Entering data

To enter data, point and click on a cell and begin typing. Pressing an arrow key will save the entry and move the data entry point to another cell. As an example, pressing the right arrow key will save the entry and move the data entry point to the right. Pressing the Escape (Esc) key will cancel data entry. Pressing the Enter key will move the data entry point down one row.

Editing data

To edit data, click on the cell to edit. Pressing any of the arrow keys will save the entry and move the data entry point to the relevant cell. Do not enter characters such as a Comma, $, or % and avoid the use of strings (words) whenever possible.

Spreadsheet and Data File Limits

The spreadsheet edit capacity is limited to 60 columns. The number of rows is controlled by system memory and speed of the computer. AcaStat was not developed for very large data files. Files larger than 100,000 observations will slow loading and analysis.

If a file has more than 60 variables, the spreadsheet will display the first 60 variables, although all the variables and observations will be available for analysis. Large files can be reduced using the Split Data module found in the Data pull-down menu. This procedure will create a new file with fewer variables. The new file will have the same number of observations and formatting as the original larger file. If a file has a large number of observations, the random sample statistical procedure can be used to create a more manageable data file.

Drag and Drop

Variables can be added to analysis list boxes by clicking on a variable, holding the mouse button, and dragging to the intended list box and releasing the mouse button. The data spreadsheet will also accept drag and drop for AcaStat data files and delimited text data files (tab and csv).

File Menu

New

Clears all data from the spreadsheet and resets the variable lists.

Open Data File

Use to open an existing AcaStat data file. AcaStat data files have the extension ".dcs". AcaStat files contain data formatting, coding of missing values, and data file notes.
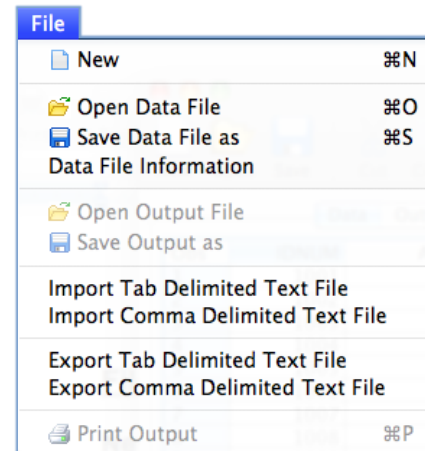
Save Data File As

Saves the contents and any formatting of the data into an AcaStat data file.

Data File Information

Displays file information and also has a section for file notes. Changes to the file notes will not be retained until the data file is saved.

Import Tab Delimited Data

Opens a text file that contains tabs between each variable so AcaStat is able to import the data into a spreadsheet format. These files normally have a filename extension of ".txt" or ".dat".

Import Comma Delimited Data

Opens a text file that contains commas between each variable so AcaStat is able to import the data into a spreadsheet format. These files normally have a filename extension of ".csv". The abbreviation "csv" represents comma separated values.
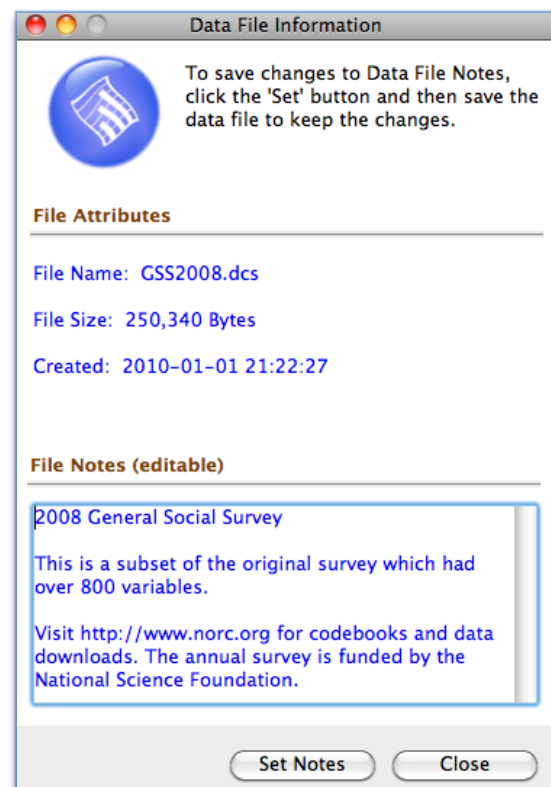
Tips on Importing Data

Import data from text data files or by pasting data copied from a word processor, spreadsheet, or database table.

Pasting Data

To paste data from a table, select and copy data in a word processor, spreadsheet, or database table. In the Data module, select "File/Import/Paste Data" or use the paste button on the toolbar. The paste procedure replaces all data in the current spreadsheet. Copied data must be consistent with the tips discussed below.
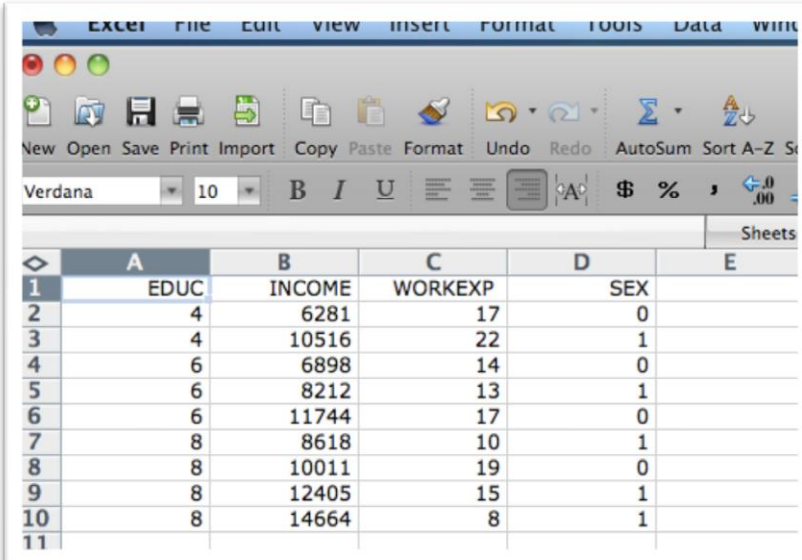
Importing Text Data Files

AcaStat assumes the first row in an imported file contains the variable names. Please see the section below on using a spreadsheet application to create compatible data files.

If the data file is large, it may take a few moments to read the data. Once read, the data will be displayed in the spreadsheet and variable names (column headings) will be visible in the variable list box. After importing data, save the file as an AcaStat system file. This will increase loading speed and allows the file to retain variable formatting. As noted earlier, use the Split Data module to reduce the size of very large data files (it is not unusual to see social science data files with over 700 variables).

Tips on using a spreadsheet application to create a csv or txt file

The first row in the spreadsheet should contain a column label that represents the data in that column (keep the labels 8 characters or less). These labels represent the variable names and identify columns when imported into the AcaStat spreadsheet (see example of spreadsheet format below). To save a delimited text file, use File/Save As and select the file type as either tab delimited or comma separated values (CSV) format.

Example of Import Error

When creating a comma separated value file, make sure all commas have been removed from the cell contents before creating the file. If they are not removed, as might be the case with currency data, AcaStat's import procedure will create an extra variable for each comma. This will cause the cell contents in the AcaStat spreadsheet to be out of alignment with the column. In the example below, the data in rows two and three do not align properly with the appropriate columns.

	AGE	SEX	INCOME	V4
1	65	M	33000	
2		23	M	
3	46000	45	F	

Other import errors may also occur if the text file does not exactly match the format specified above. The following list is not complete, but it contains common errors in importing data.

- No variable names in the first row
- Blank rows at the top of the data file
- Special embedded characters in the cells that are neither numbers nor words. This is especially common in data copied from an html table from the Internet.
- Embedded tabs in spreadsheet cells

Export Data (Tab Delimited)

Use Export Data to save the contents of the spreadsheet as a text file that contains tabs between each variable (column). These files are given the filename extension ".txt". This file can be viewed and edited in word processing or spreadsheet software and can be imported into other statistical software.

Export Data (Comma Delimited)

Use Export Data to save the contents of the spreadsheet as a text file that contains commas between each variable (column). These files are given the filename extension ".csv". This file can be viewed and edited in word processing or spreadsheet software and can be imported into other statistical software.

Data Menu

Split Large Data Files

Use Split Large Data Files to create a new data file that contains a subset of variables from the current file. This feature reduces the size of large data files to increase analysis speed and improve data editing capacity. The new file retains variable formatting.
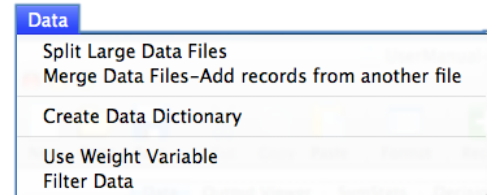
Click the Edit pull-down menu in the Data module and select the Split Large Data Files option. A list of the variables in the current data file is provided. Select variables from the current file list to include in the new file and click the right arrow button. To remove a variable from the new file list, select the variable and click the left arrow button. Click run to create the new data file.
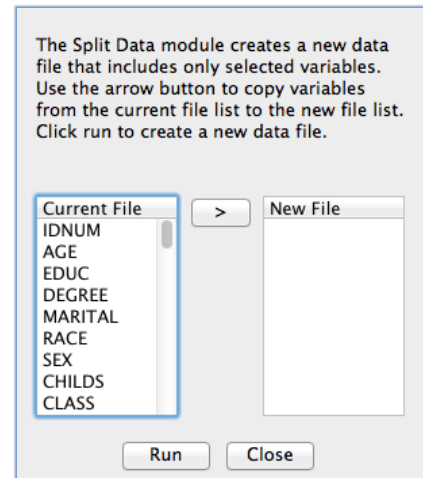
Merging Data Files to Add Records

This procedure can be used to merge two identically formatted AcaStat data files so that the observations in one file are added to the observations of another AcaStat data file. Both files must have identical formatting so that the number of variables is the same in each file and the order of the variables from left to right is identical. This procedure is especially useful if the same data formatting (variable names, labels, value labels, missing values) is used to collect new data. Examples include repeated surveys or monthly data downloads from a website or bank account.

Use the following process:

Create and save a data file with all the formatting set for variable and value labels and missing values.

Create additional AcaStat data files by entering the data or importing a delimited data file. The number and order of the variables (columns) must be identical to the original formatted data file.

To merge files,

1. Open the file that has the formatting
2. Use the menu item [Data/Merge Data Files] to select and open the second file.

The observations from the second file will be added to the first file.

Create Data Dictionary

This option creates a record of the data file variables, variable and value labels, and missing values formatting. The record is saved in the Output Viewer for saving or printing.

Use Weight Variable

Select Use Weight Variable to identify a variable for weighting observations. Weights are often used to compensate for over or under sampling and also for producing population values from a random sample. Weights can be used in the frequency and descriptive procedures. Weights will not be applied to any other procedures. Only one weight can be selected at a time.

Filter Data

Select Filter Data to exclude observations from the analysis. This module can be used to limit the analysis to a finely defined set of observations. Use up to two variables to exclude observations from subsequent analyses. The rows excluded from the analyses are marked by a bolded X. Resetting the filter removes the filter from future analysis. No data are permanently lost when using a filter.

A filter variable will be added to the spreadsheet. If the data file is saved after filtering, this filter variable will be saved and formatted with the filter specifications for later use as a control.

Run Procedure

Select Run Procedure to start an analysis. This has the same function as the Calculate button on the toolbar.

Edit Menu

Replace Values

Use Replace Values to change cell contents within one or more variables. This feature is most useful for correcting data entry errors or changing string contents to values (e.g., to change "Male" and "Female" strings so that all "Male" strings are coded the value 0 and all "Female" strings are coded the value 1). Use Undo to return the cell contents to the immediate preceding value.

This procedure can be used on one or multiple variables by holding the Control button [Mac Command button] while selecting variables.

Procedure

✓ Select the variable to search by clicking on the variable name in the list box
✓ Enter the value or string of the search parameter
✓ Enter the value to use to replace the search parameter
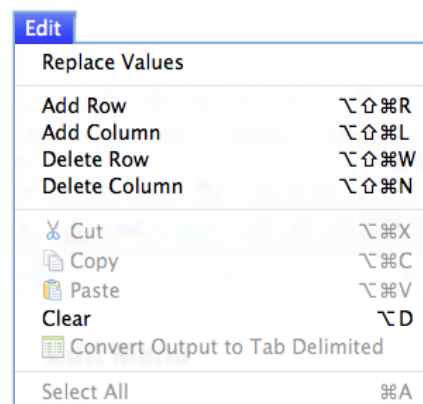
✓ Click the "Replace" button

Note: It may be helpful to conduct replace operations on duplicate variables so the original variable is not disturbed. Use the Compute procedure to duplicate a variable.

Add Row

Adds one row to the bottom of the spreadsheet.

Add Column

This adds one column to the far right of the spreadsheet up to 60 columns. In most cases, this feature will not be used. Columns are added automatically when data are entered into the next to last column.

Delete Row

Deletes the row for the currently selected cell. Delete one row at a time. The change will be complete when the data file is saved.

Delete Column

Deletes the column for the currently selected cell. Delete one column at a time. The change will be complete when the data file is saved.

Paste

When paste is used in the Data module, an import procedure places data copied from another application into the spreadsheet. Paste clears the spreadsheet before importing the data. The operation is identical to importing a tab delimited data file.

Clear

If the Data module is visible, this clears all data from the spreadsheet and resets the variable lists. If an output screen is visible, this will clear the contents and the output log.

Convert Output to Tab Delimited

If the Output module is visible, this formats the output with tabs to improve the alignment of tables copied from AcaStat for pasting into spreadsheet software. The tabs are used by spreadsheet software to align the table data properly into columns.

Variable Menu

Format Variable

To make output easier to interpret, create variable names and assign short explanatory labels to variable names and values. Formatting information will be saved with the AcaStat data file and can be changed at any time. If data are exported, the formatting information will not be exported. Click the Copy Format button to copy value and missing settings for pasting into other variables with the same format, such as survey questions with identical response options.

Variable Name

Since variable names are used as column labels and are listed in the variable list boxes, it is important to create short but meaningful variable names (8 characters or less). As an example, level of education could be named DEGREE or EDUC or EDYRS.

Variable Label

Use variable labels to create a more detailed description of the variable. The variable label is used to create more meaningful output. Using the above example, DEGREE could be labeled "Respondent's Highest Degree."

Change Decimals

Specify the number of decimals displayed (and stored) for the selected variable. This option will permanently change the variable when the file is saved.

Value Labels

Use value labels to create a more detailed description of the values used in a categorical variable. The value label is used to create more meaningful output. If education is coded 0 through 4 to represent five general categories, the values would be more meaningful if labeled 0="< High School", 1="High school", 2="Junior college", 3="Bachelor", 4="Graduate". Without labels, the output would show the values 0 through 4.

Missing Values

Identify up to four numerical values to be excluded from the statistical analysis (blank cells are always considered missing). Values that are less than (<) or greater than (>) a set value can also be set as missing.

Recoding Values

Use the values of one variable to create a new variable. This is most useful when there are several categories in a variable that can be combined into fewer meaningful groups. An example would be taking years of education and recoding the years into less than high school, high school, some college, and college graduate.

Procedure

✓ Select the variable(s) to recode by clicking on the variable name in the list box.
✓ Select the operation to perform by clicking on a radio button. For coding less than 12 years of education, select the "<" button.
✓ Enter the value to use as the basis for the recode. For the above example, enter 12.
✓ Enter a variable name for the new variable (limit to 8 characters). Example "Edufmt" for education.
✓ Enter the new variable's value to represent the old values. For this example, use 1 to represent anyone with less than 12 years of education.
✓ Click the Recode button.

A new column will be added to the spreadsheet that has the new variable name and the recoded value. Repeat the process until all values needed for the new variable have been recoded.

A module is also available to conduct multiple recodes in a single operation (up 100). Click "Set Multiple Category Ranges," select a variable, provide a name for the new variable, and then enter start and end values for each category range. The module automatically provides new values for the recodes but the values can be edited. This module should only be used with numerical data.

More than one variable can be recoded at the same time. Select and recode multiple variables by holding the Control button [Mac Command button] while selecting variables in the list box. Leave the variable name blank. AcaStat will automatically name each new variable using the original variable name and adding "-2" to the name. This option should only be used if the variables have the same range of potential values (e.g.; 1 to 10; 1,2,3,4; etc.).

Computing New Variables

Create a new variable by mathematically manipulating the values of the original variable. This is most useful when transforming a variable by squaring or taking the square root of the original values. It is also useful for applying a constant to all observations in a variable to remove negative numbers or to develop an index. The Compute procedure can be used to create an exact duplicate of a variable or create a new variable that represents the natural log of another variable's values.
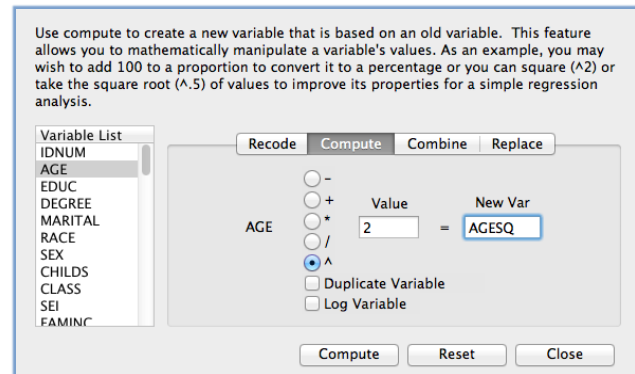
Procedure

- ✓ Select the variable to use for the compute from the list box.
- ✓ Enter a variable name for the new variable (limit to 8 characters). Example "AgeSq" for squaring age in years.
- ✓ Select the operation to perform by clicking on a radio button.
- ✓ Enter the value to use in the compute statement. Note that ^2=square and ^.5=square root.
- ✓ Click the Compute button.

A new column will be added to the spreadsheet with the variable name and the computed values for all observations.

Select and compute multiple variables by holding the Control button [Mac Command button] while selecting variables in the list box. Leave the variable name blank. AcaStat will automatically name each new variable using the original variable name and adding "-2" to the name.

Combining Two Variables

Use Combine Two Variables to create a third variable that is a mathematical combination of two other variables. This feature creates a new variable that contains the result of the computation. Repeat the operation to systematically combine several variables into one.

Procedure

- ✓ Select the first variable to use from the list box and click the Var1 button.
- ✓ Select the second variable from the list box and click the Var2 button.
- ✓ Enter a variable name for the new variable (limit to 8 characters).
- ✓ Select the operation to perform by clicking a radio button.
- ✓ Click the Combine button.
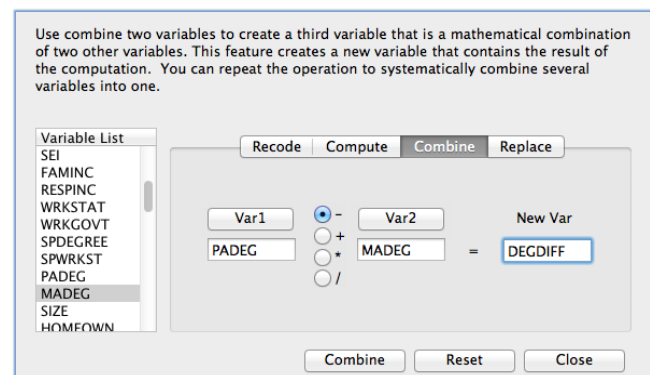
Statistical Procedures

The Statistical Procedures control panel is used to select the type of analysis and variables. An options box is displayed for each procedure.

Controls

Variable List

Use the Variable List to select the variables for analysis. As an example, to produce frequencies tables for the variables sex and race, select the sex variable in the variable list box. A [+] indicator will appear in the Analysis Variables header. To add the selected variable to the Analysis Variables list box, click the list box header or drag the variable to the Analysis Variables list box. Use the same procedure to select and place the race variable in the list box. Clicking the Calculate toolbar button will produce frequ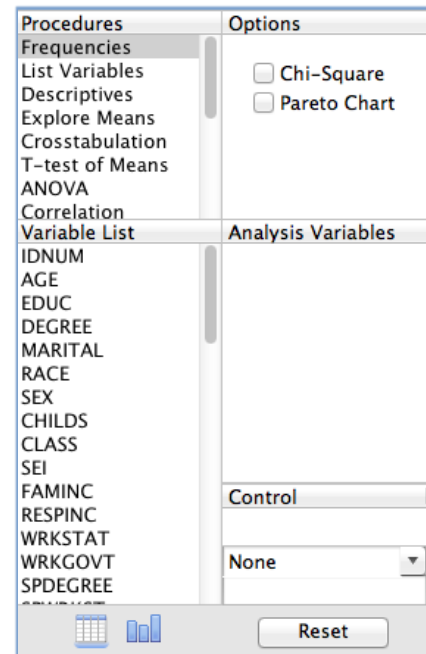encies for both variables. Output from the analyses will be displayed automatically and recorded in the Output Viewer module.

Dependent Variable

The topmost variable list box is used for all univariate analyses (Frequencies and Descriptives) and one bivariate analysis (Correlation). At least one analysis or dependent variable must be placed into the Analysis Variables list box.

Independent Variable

With the exception of Frequencies, Descriptives, and Correlation procedures, an independent variable must also be selected (another list box will appear below the Analysis Variables list box). In the case of t-tests, subgroup values must be set in the options box for the independent variable so AcaStat compares the correct groups.

Control Variable

A control can be established for all procedures. This limits the analysis to a specific subgroup. As an example, to correlate income with education for females only, locate the Sex variable name from the variable list box and drag it into the control box, click the "=" button and indicate the value used for females. This would result in a correlation procedure that excludes all males. The control should be a value (number) not a string (character).

The "Missing" option limits the analysis to records that have missing data for the control variable.

Tips for using the Control Box

If a control variable is selected without establishing a control value, the statistical procedure will run a separate analysis for each value of the control. This is handy for categorical controls such as sex and race but is not recommend for continuous controls such as age, body weight, and income unless they are first recoded into fewer groups as categorical values.

Procedures

One of the statistical procedures must be selected to run an analysis.

Frequencies

This produces a listing of all the values in a variable, the number of times the values occur in the variable, and the percent of cases represented by each value. Frequencies can be produced on numerical and string data.

List Variables

This produces a listing of observations for up to ten variables. Observations are listed in the order they appear in the data file.

Descriptives

Calculates measures of central tendency (mean, median) and variation (variance, standard deviation).

Explore Means

Explore Means produces descriptive statistics (sum, count, mean, sample standard deviation, standard error, 95% confidence interval) for each subgroup in the selected categorical variable(s). Explore requires continuous data for the analysis variable. The categorical variable can be a numeric or string variable.

Explore "Options" limits the output to selected summary statistics. This is very useful when using a categorical "By" variable. An example of the "limited" version of output is provided below.

Use the Limit option to summarize data by listing selected statistics for each category.

○ Default Output

⦿ Limit and List Output by Category:

☑ Count
☐ Sum
☑ Mean
☑ Median
☐ Standard Deviation

Cancel   OK

```
Explore Means

EDUC  by WRKSTAT


WRKSTAT                 Count         Mean           Median
-----------------------------------------------------------
Full-time                915       14.0984         14.0000
Part-time                234       13.5726         13.0000
Temp not working          33       13.8182         13.0000
Unemployed               143       12.6224         12.0000
Retired                  319       13.0721         12.0000
School                    93       13.9032         14.0000
Keeping House            234       12.1581         12.0000
Other                     65       11.7846         12.0000
```

Crosstabulation

This is a very useful procedure for representing the association between two categorical variables in a contingency table. Each cell in a contingency table represents a subgroup. For each cell, the count, row percent, column percent, and total percent are reported (default). This procedure also automatically produces a chi-square statistic. Crosstabulations are easiest to interpret when the number of rows and columns in the contingency table are 5 or less.

Crosstabulation "Options" limits the cell contents of the output. An example of the output when cell contents are limited to count and column percent is provided below.

```
Column Variable Label: Respondent Sex
Row Variable Label: Race of Respondent

        Count |
        Col % |Male       |Female    |    Total
------------------------------------------------
White         |       689|       861|     1550
              |    77.33|     74.67|
------------------------------------------------
Black         |       118|       193|      311
              |    13.24|     16.74|
------------------------------------------------
Other         |        84|        99|      183
              |     9.43|      8.59|
------------------------------------------------
              |       891|      1153|     2044
```

T-tests of means

Compare the difference between two sample means. For t-tests, the dependent variable must be a continuous variable to produce the means for analysis. The independent variable must be a categorical variable (often it will be a dichotomous variable such as sex: 0=male, 1=female). Values representing the comparison groups for the independent variable must be set in the option frame. These values are used to separate the dependent variable into two groups. Be careful to use the correct values or there may be an error reported that indicates there are not enough cases to compute the t-test. If there are more than two values in the independent variable, the other subgroups will be excluded from the analysis.

Analysis of Variance (ANOVA)

Compare the difference between two or more sample means. Like t-tests, the dependent variable must be a continuous variable to produce the means for analysis. The independent variable must be a nominal level variable (e.g., race: 1=white, 2=black, 3=Asian, 4=Hispanic. or ordinal (e.g., political ideology: 1=conservative, 2= moderate, 3=liberal). ANOVA optional output includes summary statistics for each subgroup and the Bonferroni post hoc test for multiple comparisons.

Pearson Correlation

Calculates the Pearson correlation coefficient to represent the association between two continuous variables. There are two options: create matrix or a scatterplot. The scatterplot option provides basic statistics and displays a visual representation of the association between two variables. For multiple comparisons, a correlation matrix is often the best choice.

OLS Regression

Calculates simple regression (one independent variable) or multiple regression (two or more independent variables). Ordinary least squares (OLS) regression requires a continuous dependent variable. As an option, ordinary least squares regression will add new variables to the spreadsheet representing residuals or predicted values for each observation. OLS also allows models with no constant (regression through the origin). Summary statistics for the variables included in the model are automatically included in the output.

Logistic Regression

Logistic regression requires a dichotomous (two values) dependent variable. Logistic regression will automatically produce summary statistics and, as an option, will add a new variable to the spreadsheet representing the predicted values for each observation.

Diagnostic Accuracy

Diagnostic accuracy compares a test result to a dichotomous reference standard. It is used to evaluate the sensitivity and specificity of a diagnostic test. The reference variable can only be

---

two values (usually 0 for non-disease and 1 for disease). This represents the gold standard to compare to the test results.

Each cell in the table holds counts. The test variable can be nominal, ordinal, or interval/ratio. The lowest value for the reference standard must always represent non-disease. The cutpoint (T+ >) represents the point beyond which the test is positive (disease indicated).

As an example, if possible test results range from 1 to 4, a cutpoint of 1 means a negative test result and a positive test is indicated by a value of 2, 3, or 4. A practice data file "Carotid Stenosis" is included with AcaStat.

Simulation

Two simulation procedures can be used for manipulating data in the spreadsheet. They were designed to help students develop a better understanding of random sampling. The Random Sample option creates a random sample from a larger data file. Assuming the larger data file represents the entire population, a smaller random sample helps students discover how summary sample measures such as means and proportions will vary from population parameters.

The Repeated Sampling procedure conducts multiple random samples to build a data set that reflects summary measures for each random sample. This procedure helps students test the Central Limit Theorem:

> ### Central Limit Theorem
>
> As sample size increases, the sampling distribution of means approximates a normal distribution and is close to normal at a sample size of 30.

Random Sample

This procedure is used to create a data file that is a randomly selected subset of the current spreadsheet. Enter the sample size desired. The sample size must be less than the total number of observations in the current data file. Click the Run button to replace the current data with a random sample.
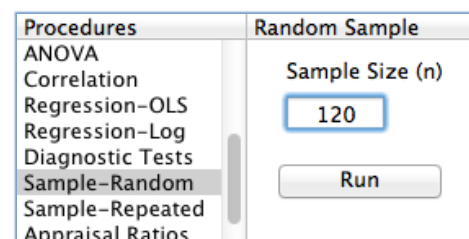
Repeated Sampling

This procedure creates repeated random samples for one variable in the current spreadsheet. For each sample (iteration), the mean and standard deviation are computed and placed into the spreadsheet. This results in a new data file where each row represents the summary statistics for one random sample.

Enter the size of the samples (n) and the number of random samples (Iterations) desired. Select a variable to conduct the random sampling on and place it in the list box. Only one variable will be used for the random sampling. Please note that the sample size must be less than the total number of observations in the current data. Click the Run button to replace the current data with a sampling distribution. When the sampling is complete, AcaStat will run the Descriptives procedure on the mean variable.

Appraisal Statistical Procedure

The Appraisal procedure is used to determine tax assessment values of real estate. Use of this procedure requires data on the market (selling) price and the assessed value of homes to compute ratios and other statistics. A practice data file "Tax Assessments" is included with AcaStat.

Charts Module

Click the chart icon below the variable list box to open the charts module. The Charts module produces graphic displays of data. The chart can be saved, printed, and copied by using the toolbar buttons or menu pull-down. The chart can also be formatted to adjust background, titles, and decimals. Only one variable at a time can be used. The control variable option available for conducting statistical procedures is also available to chart subgroups of data.
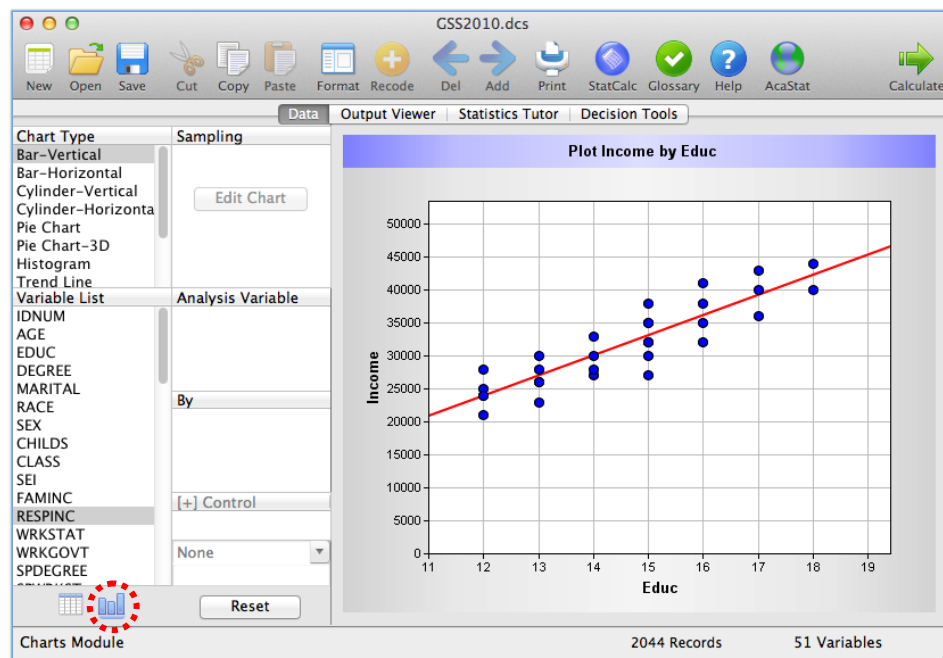
Save Chart

If using Windows and the system does not have QuickTime installed, the chart is saved as a ".bmp" file. Enter a file name and add the extension ".bmp" to ensure the file is saved as an image file. For Mac systems and Windows with QuickTime, other file type options are available such as gif and jpg.

Paste Chart

If using Windows, use Paste Special/Bitmap to paste the chart into other software such as Microsoft Word. Mac users should only need to use Paste to insert the chart into a document.

Variable Limits

Some charts only use one variable (e.g., column and pie charts), while others use two variables (scatterplot).

Editing Charts

Most features of the charts are automatic based on the data and the labeling of variables and values in the data file. However, some settings can be changed by clicking the Edit Chart button to view chart options.

Options include changing the chart title, axis labels, and the chart background. Decimal values and minimum and maximum values for the Y and X axis (currently only available for scatterplots) can also be changed.

Click OK to rebuild the chart using the new settings. Click Cancel to close without making the changes. The chart procedure will return to the default settings when a new chart is created.

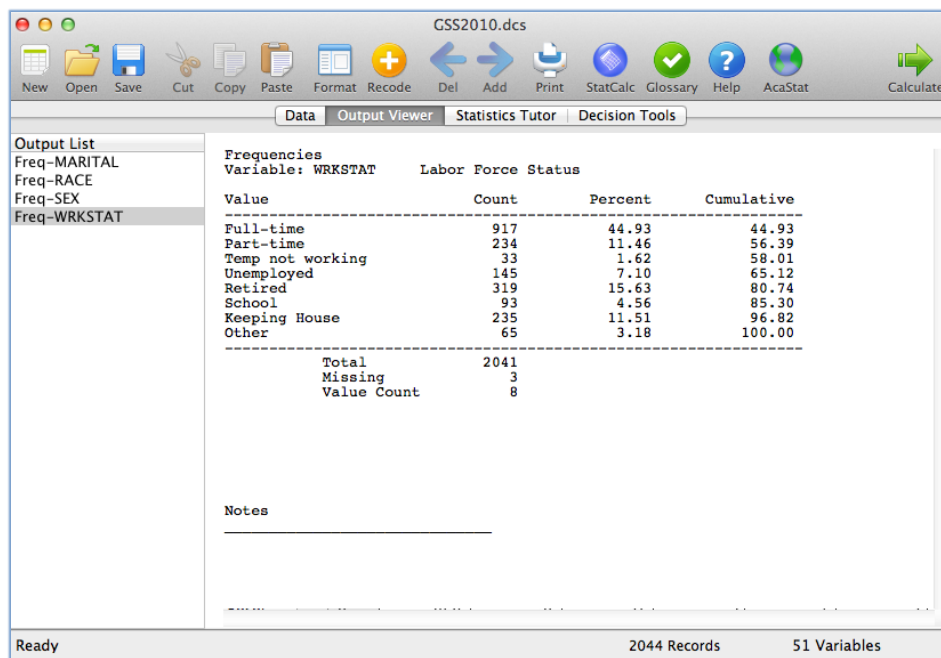## Using the "By" Variable List box

To produce a chart with group comparisons, place the grouping variable in the "By" variable list box. An example is shown below where marital status in analyzed by the sex of the respondent.
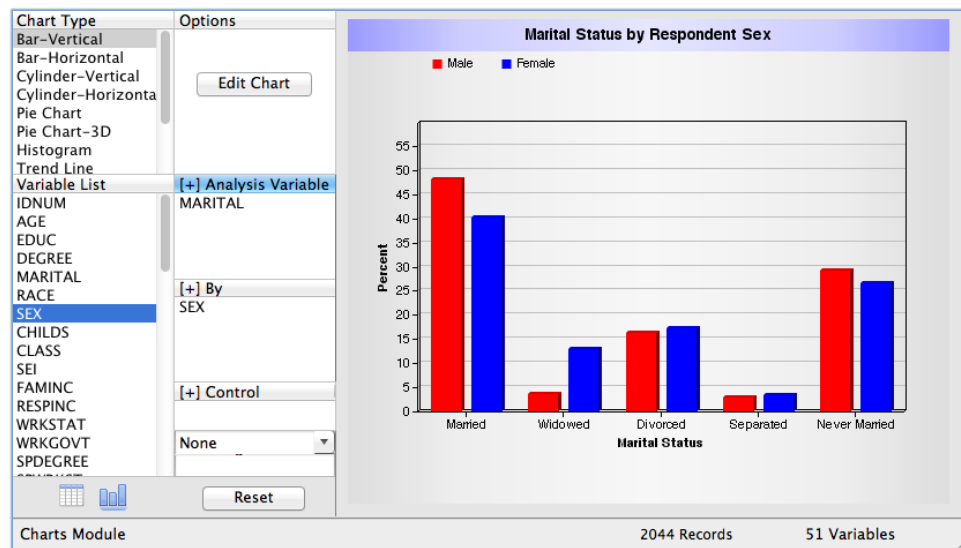
Output Viewer Module

The Output Viewer module displays the results of the analysis in a text window for editing, printing, and saving as a text file. Each time a statistical procedure is run, the results are automatically sent to the output log. Notes can be added to the output before printing or saving.

The Output Viewer uses Courier font. This is a fixed font that ensures proper alignment of the tables. When copying output to a word processor, the default font may not be Courier in the word processor. This will result in tables that are out of alignment. To correct this, highlight the output in the word processor and change the font to Courier. The font size may also need to be reduced to ensure the output fits widthwise on one page.
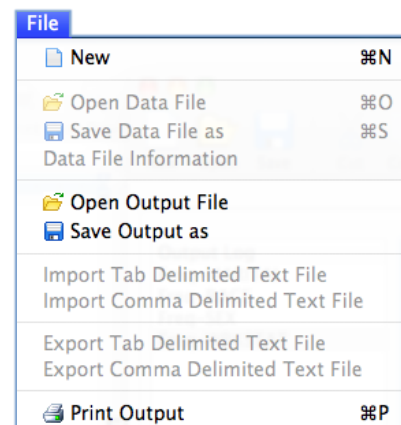

File Menu

New

Click the "New" button to erase all output.

Open Output File

This feature will load output saved from AcaStat statistical software products to include Windows AcaStat Plus and AcaStat Universal for Mac and Windows.

Save Output As

This menu item saves output as a text file with the file extension ".txt". This file can be opened in any word or spreadsheet application. Save all output or selected items in the output log. To select multiple items in the Mac version, press and hold the Command key while selecting items in the output log. For the Windows version, use the Control key for the same function.

Print Output

This menu item prints all output or selected items in the output log. To select multiple items in the Mac version, press and hold the Command key while selecting items in the output log. For the Windows version, use the Control key for the same function. By default, AcaStat prints a date and time header and a Notes section on each page printed.

Edit Menu

Cut

Use this menu item to remove selected output text and save the text to the system clipboard.

Copy

Use this menu item to copy selected output text to the system clipboard.

Paste

Use this menu item to paste text from the system clipboard into the output screen.

Clear

Use this menu item to delete all output log entries.

Select All

Selects all the text in the output screen.

Convert Output to Tab Delimited

Use this option to convert output to improve copying and pasting tables into a spreadsheet. This option inserts tabs into the output so that pasted output aligns properly in the columns of a spreadsheet.
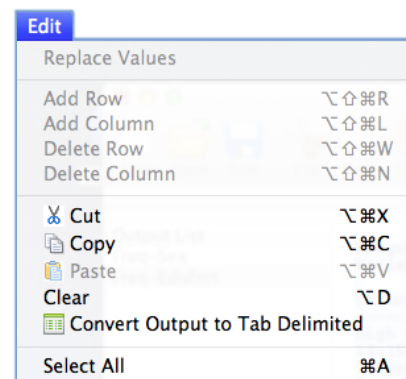
Use the Edit menu item Convert Output to Tab Delimited to convert the text. The operation will automatically select the converted text. Click the copy button and open a spreadsheet program, select a cell in the spreadsheet, and click the paste button in the spreadsheet. Check the results carefully to ensure the table is formatted and aligned properly.

Statistics Tutor

A Statistics Tutor is provided for reviewing basic statistics. Use the pull-down menu to select a lesson. When using the tutorial, the calculation output will be displayed in the information screen. More detailed output is also logged in the Output Viewer module. The Tutor displays black instructional text and blue questions and additional information. Click on the blue text to check answers to questions or for more information regarding the subject being covered. The evaluation version limits the tutorial to the first lesson (descriptive statistics).
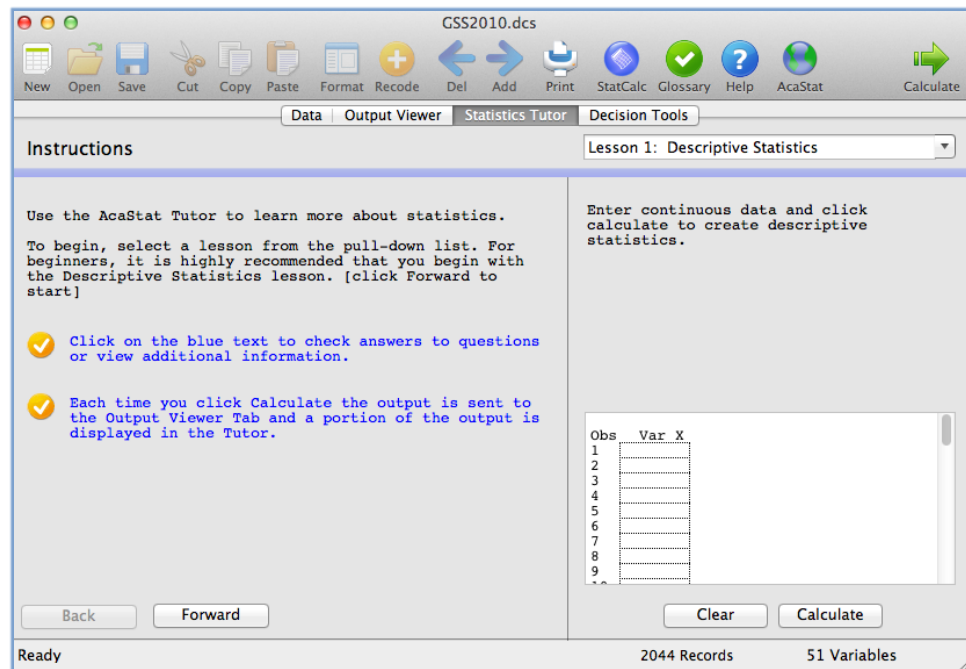
Tutorial Review Order

For many users, it may be preferable to begin with descriptive statistics followed by standardized scores, confidence intervals, z-tests, and t-tests. The tutorial cannot replace a comprehensive course on statistics but it can be used to review and practice statistical concepts and techniques.

Decision Tools Module

The Decision Tools module provides four tools for calculating constant dollars and experimenting with decision tables, queuing theory (waiting lines), and price elasticity of demand. An example is provided for each tool. The results are displayed in the Decision Tool tab and are added to the output log. A good reference text for most of these tools is *Fundamentals of Management Science*, by Efraim Turban & Jack R. Meredith.

Decision Tables

Decision tables quantify in table form the impact of decision options under uncertainty. Tables will typically contain the following elements:

1. Alternative courses of action
2. States of nature
3. Probabilities of states of nature (must sum to 1)
4. The payoffs

The example for the Decision Table is displayed below.

Queuing Theory

Queuing theory is used to evaluate the performance of service systems where waiting lines may form and customers arrive at random times (i.e., license bureau, tax assessors office, emergency room). To evaluate a waiting line, data must be provided for the average arrival rate, average service rate, and the number of servicers.

Constant Dollars

This tool uses the urban consumer price index (CPI-U) for the United States to adjust for inflation. The module will create and apply an inflation and deflation estimate. Another price index can be used by overriding the default CPI by entering "*" followed by a CPI.

Price Elasticity of Demand

This tool uses the change in demand that occurs when a price is changed to estimate price elasticity. Given the estimate of elasticity, the tool will attempt to find a breakeven price through iterations of price. Both fixed and variable costs can be included in the model.

StatCalc Module

StatCalc is a separate module that simplifies common statistical calculations with 18 procedures. Open StatCalc from the toolbar or the tools menu. Use StatCalc to create and compare summary statistics, develop confidence intervals, and verify hand calculations. Enter data to create descriptive, correlation, and regression statistics. This special version of StatCalc allows 200 observations. The stand-alone version of StatCalc can be purchased separately. It allows 5000 observations. The AcaStat Data spreadsheet should be used for large data entry projects.
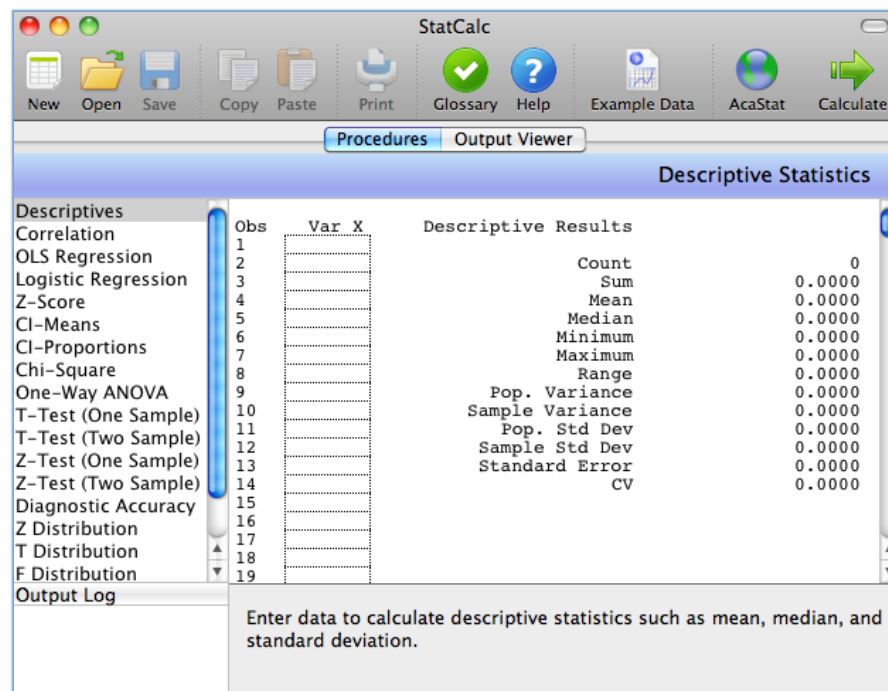
StatCalc includes example data for each module. All output is stored in the output log in both StatCalc and AcaStat to view, save, and print results.

Entering Data

To enter data, point and click on a cell and begin typing. Pressing the right, left, up, or down arrow key will save the entry and move the data entry point to the next cell. Pressing the Enter (Return) key will move the data entry point to the next row.

Importing Data Files

To import a text file into StatCalc, click File/Import Data to select the data file. This feature is only available for Descriptives, Correlation, OLS Regression and Logistic Regression. The data file must be a tab or comma delimited text file with variable names in the first row of data. Create or export these text files from most software packages to include word processors, spreadsheets, databases, and statistical packages. The import is limited to 200 observations.

Exporting Data

Click File/Export Data or the save toolbar button to save the data in the spreadsheet as a tab or comma separated (delimited) file. This feature is only available for Descriptives, Correlation, OLS Regression and Logistic Regression. Exported files can be easily be viewed in other software packages.

If the exported file is opened in spreadsheet software, the first row will list the variable names and the remaining rows will contain the data (one row for each record). Each column represents the data for one variable.

Diagnostic Accuracy

Diagnostic accuracy compares a test result to a dichotomous reference standard. It is used to evaluate the sensitivity and specificity of a diagnostic test. The reference and test variables can only be two values (usually 0 for non-disease and 1 for disease). The reference variable represents the gold standard to compare to the test results.

Logistic Regression

Calculates simple regression (one independent variable) or multiple regression (two or more independent variables). Logistic regression requires a dichotomous (0 and 1) dependent variable to develop a model that predicts the occurrence of 1.

OLS Regression

Calculates simple regression (one independent variable) or multiple regression (two or more independent variables). OLS regression requires an interval/ratio dependent variable.

Descriptives

Calculates measures of central tendency (mean, median) and variation (variance, standard deviation). Descriptives requires continuous data. Enter the data and click calculate when done.

Chi-Square

This is a very useful procedure for representing the association between two categorical variables via a contingency table. Each cell in a contingency table represents a subgroup. All the rows and columns do not need to be used but always start with the top leftmost cell. Enter counts in each cell and StatCalc with calculate for each cell the row percent, column percent, and total percent.

This procedure also automatically produces several contingency table statistics to include the chi-square statistic. Crosstabulations are easiest to interpret when the number of rows and columns in the contingency table are 5 or less.

Correlation

Calculates the Pearson correlation coefficient to represent the association between two interval/ratio level variables. Enter the data as indicated and click Calculate when done.

Regression

Calculates simple regression (one independent variable) or multiple regression (two or more independent variables). Ordinary least squares regression requires interval/ratio level dependent variables and logistic regression requires a dichotomous dependent variable. Enter the data as indicated and click Calculate when done.

T-test One Sample

Used to test for the difference between one population mean (mu) and one sample mean. Enter the population mean, sample mean, sample standard deviation and the number of observations in the sample.

T-test Two Sample

Used to test for the difference between two sample means. For t-tests, the dependent variable must be an interval/ratio level variable to produce the means for analysis. The independent variable must be a nominal or ordinal level variable (often it will be a dichotomous variable such as sex: 1=male, 2=female).

Z-test One Sample

Used to test for the difference between one population proportion and one sample proportion. Enter the population proportion, sample proportion, and the number of observations in the sample.

Z-test Two Sample

Used to test for the difference between two sample proportions. For two sample z-tests, both dependent and independent variables must be nominal level variables.

Z Distribution

Enter a z score to calculate the proportion of the probability distribution at or beyond the score entered. This procedure is most useful for evaluating the statistical significance of a z-test statistic calculated manually.

T Distribution

Enter a t score and degrees of freedom to calculate the proportion of the probability distribution at or beyond the score entered. This procedure is most useful for evaluating the statistical significance of a t-test statistic calculated manually.

F Distribution

Enter a f ratio score and numerator and denominator degrees of freedom (consult a statistical reference) to calculate the proportion of the probability distribution at or above the score entered. This procedure is most useful for evaluating the statistical significance of an f test statistic calculated manually.

Chi Square Distribution

Enter a chi-square score and degrees of freedom to calculate the proportion of the probability distribution at or above the score entered. This procedure is most useful for evaluating the statistical significance of a chi-square test statistic calculated manually.

One-Way ANOVA

Used to test for the difference between two or more sample means. The dependent variable must be an interval/ratio level variable to produce the means for analysis. The independent variable must be a nominal level variable (e.g., race: 1=white, 2=black, 3=Asian, 4=Hispanic. or ordinal (e.g., political ideology: 1=conservative, 2= moderate, 3=liberal). All the rows and columns provided do not have to be used but always start with the top leftmost cell.

Output

The Output Viewer module displays the results of the analysis in a text viewer for editing, printing, and saving as a text file. Each time a statistical procedure is run, the results are automatically sent to the output log. Each subsequent transfer of results is appended to the output log creating a complete record of the analyses.

Click the Clear button to erase all output or select specific items and press the Delete key to remove the item. Multiple items can be selected by simultaneously pressing the Control button (Command button for Mac systems) while clicking items with the mouse pointer.

Printing Output

All or selected output results can be printed. Select specific output results by holding the Control button while selecting output from the list box.

Save Output

The contents of the Output Viewer module can be saved as a text file. Similar to printing, all or selected output can be saved.

Copying/Pasting

The Output Viewer uses Courier font. Some word processing software will format the pasted text to match the default font setting. This will cause the pasted text to become misaligned. When copying and pasting the results into another document ensure the font of the output is set to Courier New or the tables may not align properly.

Other Menu Items and AcaStat Features

Data Examples Menu

Several practice data files are available for use with AcaStat.
The data files include variable and value formatting. The files
can be saved for manipulation and use as instructional material.

The following data sets are based on actual research surveys.

General Social Surveys 2010, 2008 and 1993
AFQT
City Managers
College Admissions
Student Workbook
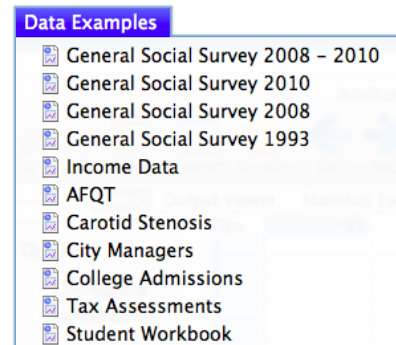
The following data sets are fictional.

Carotid Stenosis
Tax Assessments
Income Data

Documentation

Use this pull-down menu to open application documentation and
statistics instructional material.
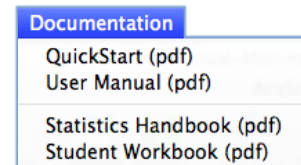
Help Menu

Help

Click Help to open the help module. The contents can
be sorted or searched by topic.
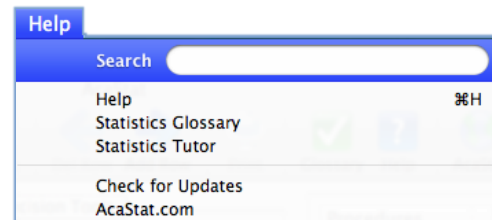
Statistics Glossary

The glossary contains over 200 terms and definitions. The search button will find matching
records and sort the list box so all topics matching the search are listed at the top in bold.
Pressing the Enter key after entering a search term will also initiate the search. Press Reset
before conducting another search.

Check for Updates

The Help menu contains an option to check for updates. AcaStat connects to the Internet and automatically compares the installed version with the version of the update. If a newer update is available, the appropriate installation file may be downloaded to the computer desktop.

The update option is only available for licensed users.

Applied Statistics Handbook

The Applied Statistics Handbook contains formulas, definitions, and examples along with annotated output from statistical software. It is designed to be a quick reference resource (not a complete statistics text).

If you have a licensed version of AcaStat, the Handbook can be opened via the AcaStat Documentation menu. All AcaStat manuals are copyright protected.

File Name: Statbook.pdf
Availability: AcaStat Licensed Edition

Student Workbook

The Student Workbook contains over 20 lessons designed to introduce basic statistical concepts while also demonstrating the functionality of AcaStat.

The Student Workbook can be opened via the AcaStat Documentation menu.

File Name: Workbook.pdf
Availability: AcaStat CD, Website http://www.acastat.com/pdfdocs.htm

Review of Features

AcaStat is designed to provide students, teachers, researchers, and managers with an inexpensive and easy to use tool for data exploration. It includes three data analysis modules (Data, StatCalc, Decision Tools modules) and three support modules containing analysis output (Output Viewer module), statistical terms and definitions (Glossary module), and a tutorial. Special pricing and site licenses are available for instructors adopting AcaStat for course work. Printable manuals include a Student Workbook with over 20 lessons, an Applied Statistics Handbook, and an AcaStat User Manual.

Data Spreadsheet

Analyze raw data with this spreadsheet design. Produce descriptive and inferential statistics, and creates crosstabulations. Includes the features listed below.

Feature	Details
Spreadsheet	create data files, maximum edit capacity 60 columns
Import Data	import comma and tab delimited text files, paste tables from other software
Data Formatting	create variable and value labels, set missing values
Variable Recode	compute, combine, and recode variables
Paste Data	copy and paste data from spreadsheets, documents, and database tables
Export Data	export comma and tab delimited text files
Output	edit output, save as text file, print, copy, paste
Statistics	
Frequencies	frequencies (count, percent, total)
Descriptives	mean, median, variance, standard deviation
t-test of means	
One-way ANOVA	summary statistics and Bonferroni post hoc analysis
Chi-square	Cramer's V, Pearson C, Lambda, Kappa, Fisher's, odds ratio
Pearson correlation	correlation matrix and scatterplot
Simple regression	OLS
Multiple regression	OLS and Logistic
Sampling	Random sample and repeated random sampling
Explore Means	
Diagnostic Accuracy	Sensitivity, specificity, likelihood ratios, predictive values
Appraisal	Ratios and summary statistics

Tutor

Tutor provides summary instructions and interpretations of statistical output.

Feature	Details
Statistical Tests	one and two sample t-test of means one-way ANOVA (6 sample) one and two sample z-test of proportions chi-square (6x6 table)
Raw Data Analysis	Descriptive statistics Correlation OLS Simple and Multiple Regression
Margin of Error-Mean	90%-95%-99% confidence intervals, margin of error, upper limits, lower limits
Margin of Error-Proportions	90%-95%-99% confidence intervals, margin of error, upper limits, lower limits
Standardized Scores	z-score, graphic representation on normal distribution
Diagnostic Accuracy	sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, negative likelihood ratio

System Requirements (recommended minimum)

Mac OS X v. 10.2 or later
Windows 98 or later

AcaStat Software
43584 Merchant Mill Terrace
Leesburg, VA 20176 USA
support@acastat.com
http://www.acastat.com